ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 426 (2007) 855–858
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Book review

Positive Definite Matrices, Rajendra Bhatia, Princeton University Press (2007)

This book which is of high standard is devoted to in depth studies of positive (definite) matrices and related objects and is considered as an extension of the author's earlier highly appreciated monograph "Matrix Analysis" (Springer, 1996). Though the author says that the book is oriented towards those interested in linear algebra and matrix analysis, operator theorists and theoretical engineers can benefit much from this book through learning beautiful applications of various methods to matrices. Many new results in the book are taken from the contributions of the author himself. The book is a welcomed source for research in future. Graduate students and even researchers can benefit from many exercises of a high level with useful hints.

Let me survey the main topics of each chapter. Chapter 1 (Positive Matrices) is a quick review of some of the basic properties of positive (definite) matrices. Since $\mathbb{H}_n$, the space of $n \times n$ Hermitian matrices, is provided with the order relation induced by the cone $\mathbb{P}_n$ of positive matrices, one can speak about monotonicity and convexity (and concavity) of a map from $\mathbb{P}_n$ or $\mathbb{H}_n$ to $\mathbb{H}_n$. As important examples, it is shown that the map $A \longmapsto A^r$ on $\mathbb{P}_n$ is convex or monotone according $1 \leqslant r \leqslant 2$ or $0 < r \leqslant 1$ (Löwner Theorem). The norm inequalities by the author for such a map are included: for the spectral (or operator) norm $\|\cdot\|$

$$\|A^r + B^r\| \leqslant \|(A+B)^r\| (1 \leqslant r < \infty) \quad \text{and} \quad \|A^r + B^r\| \geqslant \|(A+B)^r\| \quad (0 \leqslant r \leqslant 1).$$

In Chapter 2 (Positive Linear Maps), linear maps Φ from $\mathbb{M}_n$ to $\mathbb{M}_k$, which transform $\mathbb{P}_n$ to $\mathbb{P}_k$, are studied. Such a map Φ is said to be positive and is monotone on $\mathbb{H}_n$. It is said to be unital if $\Phi(I) = I$.

Many basic inequalities for a unital positive linear map Φ (due to Kadison, Choi and others) are included. For instance,

$$\Phi(A^*)\Phi(A) \leqslant \Phi(A^*A) \text{ (normal } A \in \mathbb{M}_n) \quad \text{and} \quad \Phi(A)^{-1} \leqslant \Phi(A^{-1})(A \in \mathbb{P}_n).$$

An inequality of the Kantorvich type is also included: with $\alpha \leqslant \beta$,

$$\alpha I \leqslant A \leqslant \beta I \Longrightarrow \Phi(A^{-1}) \leqslant \frac{(\alpha+\beta)^2}{4\alpha\beta}\Phi(A)^{-1}.$$

Useful is the fact that the (mapping) norm of a positive linear map Φ, considered as $(\mathbb{M}_n, \|\cdot\|) \longmapsto (\mathbb{M}_k, \|\cdot\|)$, is attained at the identity matrix : $\|\Phi\| = \|\Phi(I)\|$ (Russo-Dye Theorem).

The notion of positivity is extended to a linear map from a subspace $\mathscr{S}$ of $\mathbb{M}_n$, which contains the identity I and is closed for adjoint formation. Such a subspace is called an operator system. For a linear functional φ on an operator system its positivity and the norm condition $\varphi(I) = \|\varphi\|$ are equivalent. As an application, a positive version of the Hahn-Banach extension theorem is

doi:10.1016/j.laa.2007.05.013

included: every positive linear functional on an operator system can be extended to a positive linear functional on $\mathbb{M}_n$ (Krein Theorem).

At the end of Chapter 2, the notion of operator monotonicity of a real valued function $f(t)$ on the (open or closed) half line of $\mathbb{R}$ is introduced as the monotonicity of the map $A \longmapsto f(A)$ on $\mathbb{P}_n$ (without respect to n). Similarly, operator convexity and operator concavity of such a function are defined. The notion of unitary invariance of a norm $|||\cdot|||$ on $\mathbb{M}_n$ is introduced.

In Chapter 3 (Completely Positive Maps), observation, not very familiar to matrix analysts, is developed. A linear map Φ from an operator system $\mathscr{S} \subset \mathbb{M}_n$ to $\mathbb{M}_k$ gives rise to a linear map Φ_m from $\mathbb{M}_m(\mathscr{S})$ to $\mathbb{M}_m(\mathbb{M}_k)$ by $\Phi_m([[A_{ij}]]) = [[\Phi(A_{ij}]](A_{ij} \in \mathscr{S})$. The space $M_m(\mathbb{M}_n)$ is canonically identified with $\mathbb{M}_{mn}$ and $\mathbb{M}_m(\mathscr{S})$ becomes its operator system. When Φ_m is positive for all $m = 1, 2, \ldots$, the map Φ is said to be completely positive. A finite dimensional version of the basic structure theorem for a completely positive map is included. For a completely positive map ϕ from $\mathbb{M}_n$ to $\mathbb{M}_k$ there exist an algebraic *-representation $\Pi : \mathbb{M}_n \longmapsto \mathbb{M}_{n^2k}$ and an $n^2k \times k$ matrix V such that $\|V\|^2 = \|\Phi(I)\|$ and $\Phi(A) = V^*\Pi(A)V(A \in \mathbb{M}_n)$ (Stinespring Dilation Theorem). The requirement for complete positivity seems too much of a burden. But there is a simple criterion (Theorem of Choi and Kraus) that Φ from $\mathbb{M}_n$ is completely positive if $[[\Phi(E_{ij})]] \geqslant 0$, where E_{ij} $(i, j = 1, \ldots, n)$ are the matrix units of $\mathbb{M}_n$. In this case there exist $n \times k$ matrices V_j $(j = 1, 2, \ldots, nk)$ such that $\Phi(A) = \sum_{j=1}^{nk} V_j^* A V_j (A \in \mathbb{M}_n)$. Therefore to see complete positivity of Φ one should check the positivity of Φ_n only.

The most fundamental fact is included, which is an extension of the Krein theorem. A completely positive map from an operator system $\mathscr{S} \subset \mathbb{M}_n$ to $\mathbb{M}_k$ can be extended to a completely positive map from $\mathbb{M}_n$ to $\mathbb{M}_k$ (Arveson Extension Theorem). When combined with the Stinespring dilation theorem, the Arveson extension theorem can give very useful information to the matrix theory. As example two results are included. One is concerned with the norm of the Schur multiplication map. Given $T \in \mathbb{M}_n$, consider the map $S_T(A) \equiv T \circ A$ (Schur (or entrywise) multiplication). It is shown that $\|S_T\| \leqslant 1$ if and only if there exist $R_j \in \mathbb{P}_n$ with $I \bigcirc R_j \leqslant I (j = 1, 2)$ such that $\begin{bmatrix} R_1 & T \\ T^* & R_2 \end{bmatrix} \geqslant 0$ (Haagerup Theorem). The other is for numerical radius $w(A) \equiv \sup\{|\langle x, Ax\rangle| : \|x\| = 1\}$ of $A \in \mathbb{M}_n$. It is shown that $w(A) \leqslant 1$ if and only if there is $H \in \mathbb{H}_n$ such that $\begin{bmatrix} I+H & A \\ A^* & I-H \end{bmatrix} \geqslant 0$ (Ando Theorem). An extension of an inequality of the Kantorovic type is also included.

In Chapter 4 (Matrix Means), generalizing the case of a pair of positive numbers the notion of (matrix) mean is introduced. A mean M is a binary operation for a pair $(A, B) \subset \mathbb{P}_n$ which satisfies the following conditions:

(i) $M(A, B) \in \mathbb{P}_n$,
(ii) $A \leqslant B \Longrightarrow A \leqslant M(A, B) \leqslant B$,
(iii) $M(A, B) = M(B, A)$,
(iv) $M(A, B)$ is monotone in each A, B,
(v) $M(X^*AX, X^*BX) = X^*M(A, B)X$ for invertible X,
(vi) $M(A, B)$ is continuous in A, B.

Obvious examples are the arithmetic mean $\frac{A+B}{2}$ and the harmonic mean $\left(\frac{A^{-1}+B^{-1}}{2}\right)^{-1}$. The most interesting and important is the geometric mean $A\sharp B$, defined as

$$A\sharp B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.$$

Monotonicity of the map $A \longmapsto A^{1/2}$ is essential for condition (iv). Various characterizations for the geometric mean $A\sharp B$ are mentioned. For instance, $A\sharp B$ is a unique positive solution of the matrix equation $XA^{-1}X = B$. This reveals a role of the geometric mean in the theory of matrix (algebraic) Riccati equations. Another characterization is

$$A\sharp B = \max\left\{X \geqslant 0;\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geqslant 0\right\}.$$

The matrix arithmetic–geometric mean inequality and the matrix geometric-harmonic mean inequality are immediate.

The Furuta's extension of monotonicity of the map $A \longmapsto A^t (0 < t \leqslant 1)$ is included:

$$A \geqslant B > 0 \longmapsto (B^r A^p B^r)^{1/q} \geqslant B^{(p+2r)/q} \quad \text{for } p \geqslant 0, r \geqslant 0, q \geqslant 1, q \geqslant \frac{p+2r}{1+2r}.$$

Some part of Chapter 4 is devoted to concave and convex inequalities related to quantum entropies. It starts from the fact that for $0 < r < 1$ the map $f(A, B) \equiv A^r \otimes B^{1-r}$ is jointly concave and monotone on the pairs of positive matrices (A, B) (a variant of Lieb Theorem). As a consequence for any $X \in \mathbb{M}_n$ the map $(A, B) \longmapsto \operatorname{tr}(X^*A^rXB^{1-r})$ is jointly concave on the pairs $(A, B) \subset \mathbb{P}_n$. A positive matrix with trace 1 is called a density matrix. The entropy $S(A)$ of a density matrix A is by definition $S(A) = -\operatorname{tr}(A \log A)$. The relative entropy of density matrices A, B is defined as $S(A|B) = \operatorname{tr}A(\log A - \log B)$, which is non-negative quantity. It follows from the Lieb Theorem that the relative entropy $S(A|B)$ is jointly convex on the pairs of (A, B).

It is pointed out that for a positive operator monotone $f(t)$ such that $f(1) = 1$ and $xf(x^{-1}) = f(x)$ the map $M(A, B) = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}$ gives rise to a mean satisfying (i) to (vi).

Obviously the arithmetic mean, the geometric mean and the harmonic mean are given by the functions $\frac{1+x}{2}$, $\sqrt{x}$, and $\frac{2x}{1+x}$ respectively. The mean corresponding to the function $\frac{x^r+x^{1-r}}{2}$ $(0 \leqslant r \leqslant 1)$ is named the Heinz mean while the one corresponding to $\frac{x-1}{\log x}$ is named the logarithmic mean. The importance of the logarithmic mean is explained in connection with heat flow.

Chapter 5 (Positive Definite Functions) is closely related to harmonic analysis. Recall that a (doubly infinite) sequence of complex numbers $\{a_n : n \in \mathbb{Z}\}$ is said to be positive definite if for every positive integer N the $N \times N$ matrix $[[a_{i-j}]]$ is positive.

Correspondingly a complex valued function φ on $\mathbb{R}$ is positive definite if for every positive integer N and every choice of real numbers $x_0, x_1, \ldots, x_{N-1}$ the $N \times N$ matrix $[[\varphi(x_i - x_j)]]$ is positive.

On the basis that the set of positive definite functions on the real line (or positive definite sequences) is closed for product, addition and multiplication by positive numbers the author presents many useful examples of positive definite functions and sequences. Also a simple proof is included for the Bochner theorem (and the Herglotz theorem) that a function (and a sequence) is positive definite if and only if it is the Fourier transform (and the Fourier coefficients) of a positive measure on the real line (and on the unit circle). The positivity of the matrix induced by a positive definite function is used to derive some norm inequalities related to matrix means.

In Chapter 6 (Geometry of Positive Matrices) the manifold $\mathbb{P}_n$ is observed as a nice object of Riemanian geometry. Since the tangent space at each point is identified with $\mathbb{H}_n$, the positive quadratic form $\langle\cdot,\cdot\rangle_A$ in the tangent space at A is defined as $\langle H, K\rangle_A = \operatorname{tr}(A^{-1}HA^{-1}K)$. Then the distance $\delta_2(A, B)$ between A and B is invariant for the congruent transformation $\Gamma_X(A) = X^*AX$ with invertible X. This makes it possible to give an explicit representation of the distance as $\delta_2(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|_2$. Here $\|X\|_2$ is the Frobenius (or Hilbert–Schmidt) norm

$\|X\|_2 = \mathrm{tr}(X^*X)^{1/2}$. It turns out that a unique geodesic curve $\gamma(t)$ connecting A and B is given by

$$\gamma(t) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2} \quad (0 \leqslant t \leqslant 1).$$

The geometric mean $A\sharp B$ is characterized as the midpoint of the geodesic connecting A and B.

From the definition of the metric δ_2 it follows that the exponential function from $(\mathbb{H}_n, \|\cdot\|_2)$ to $(\mathbb{P}_n, \delta_2)$ is distance-increasing $\delta_2(e^H, e^K) \geqslant \|H-K\|_2 (H, K \in \mathbb{H}_n)$. It follows from this that the parallelogram law in the $(\mathbb{H}_n, \|\cdot\|_2)$

$$\left\|\frac{1}{2}(H+K) - L\right\|_2^2 = \frac{\|H-L\|_2^2 + \|K-L\|_2^2}{2} - \frac{\|H-K\|_2^2}{4}$$

is transferred to the semiparallelogram law in $(\mathbb{P}_n, \delta_2)$ as

$$\delta_2^2(A\sharp B, C) \leqslant \frac{\delta_2^2(A,C) + \delta_2^2(B,C)}{2} - \frac{\delta_2^2(A,B)}{4}.$$

This is used to show that the geometric mean $A\sharp B$ is characterized as a unique point at which the function $f(X) = \delta_2^2(A, X) + \delta_2^2(B, X)$ attains the minimum.

Inspired by this fact, the author proposes to define a geometric mean of a triple $(A_1, A_2, A_3) \subset \mathbb{P}_n$ as unique point X_0 at which the function $f(X) = \sum_{i=1}^3 \delta_2^2(A_i, X)$ attains the minimum, and shows that such X_0 is a unique positive solution of the matrix equation $\sum_{i=1}^3 X^{-1}\log(XA_i^{-1}) = 0$. This geometric mean is considered as the center of mass of a triple (A_1, A_2, A_3).

Another candidate of a geometric mean of a triple (A_1, A_2, A_3), inspired by the characterization of the center of a plane triangle as the intersection of a sequence of nested triangles, is also mentioned.

Define the sequence of the triples $(A_1^{(m)}, A_2^{(m)}, A_3^{(m)})$ as $A_i^{(0)} = A_i (i = 1, 2, 3)$ and

$$A_1^{(m+1)} = A_1^{(m)}\sharp A_2^{(m)}, \quad A_2^{(m+1)} = A_2^{(m)}\sharp A_3^{(m)}, \quad \text{and} \quad A_3^{(m+1)} = A_3^{(m)}\sharp A_1^{(m)}.$$

Then $A_i^{(m)}$ converges, as $m \to \infty$, to one and the same limit for $i = 1, 2, 3$. This limit is a candidate of a geometric mean. The above two candidates of a geometric mean do not coincide in general.

In the final place it is mentioned briefly what happens if $\mathbb{P}_n$ is provided with a Finsler metric $\delta_{|||\cdot|||}(A, B) = |||\log(A^{-1/2}BA^{-1/2})|||$, based on a unitarily invariant norm $|||\cdot|||$. For this norm also the inequality $|||H-K||| \leqslant \delta_{|||\cdot|||}(e^H, e^K)$ is valid, which implies $|||e^{H+K}||| \leqslant |||e^{H/2}e^K e^{H/2}|||$.

Specialised to the cases of the trace norm and the spectral norm this inequality gives the well-known inequalities in physics (Golden–Thompson inequality and Segal inequality).

T. Ando
Shiroishi-ku, Hongo-dori 9,
Minami 4-10-805,
Sapporo 0003-0024, Japan
E-mail address: ando@es.hokudai.ac.jp

Available online 24 May 2007